

Третьякова Т.Г.

ТЕСТИРОВАНИЕ КОМПЛЕКСНЫХ СВОЙСТВ НА ОДНОМЕРНОЙ ШКАЛЕ ТЕСТАМИ НОВОГО ТИПА

tretjakova@newmail.ru

Новосибирский государственный педагогический университет

г. Новосибирск

Зачастую встречается необходимость тестирования свойств, являющихся комплексными. Например, тестовые задания по физике неизбежно предполагают наличие у испытуемых определенных математических знаний. Здесь возникают два подхода. Первый подход предполагает существенное уменьшение требований к необходимым для выполнения тестовых заданий по физике математических знаний и поддержка этих требований к заданиям на минимальном уровне во всем диапазоне трудностей заданий. Формально при этом повышается валидность тестовых результатов, однако нарушается другое требование к тестовым заданиям – требование репрезентативности, поскольку они должны представлять собой репрезентативную выборку из генеральной совокупности всех тестовых заданий по физике, где какая-либо связь между физическим содержанием задачи и ее математической сложностью отсутствует.

Второй подход предполагает наличие прямой связи между физической и математической трудностью задачи. Оправдан такой подход изучением обеих учебных дисциплин в рамках единого учебного плана (например, в средней школе), и на практике реализуется отставанием математического уровня трудностей задания на фиксированный промежуток от физического уровня трудности задания.

Однако существует и третий подход. При выделении n_ϕ уровней физической трудности заданий и n_m уровней математической трудности в тесте необходимо применить $n_m = n_\phi n_m$ заданий со всеми комбинаторно возможными сочетаниями уровней физической и математической трудности. В этом случае тест перестает быть тестом по физике и начинает измерять некоторое комплексное свойство (знание физики и математики) на одномерной шкале.

В данном докладе мы не касаемся целесообразности введения в практику тестирования тестов подобного рода именно по физике и математике, поскольку подобные вопросы решаются, как правило, исходя из сложившейся в рамках действующих учебных планов традиций.

Однако в практике тестирования таких предметов, как иностранный язык, тестирование отдельных аспектов владения языком зачастую не предусмотрено, а общая оценка уровня владения языком необходимо должна включать в себя отдельные аспекты владения языком, и быть тем не менее одномерной. Возникает проблема шкалирования подобных тестов в рамках привычной классической теории тестирования, где измеряемое свойство оценивается исходным (первичным, сырым) тестовым баллом. Однако подобное правомерно лишь для гомогенного теста, измеряющего однородное свойство. Использование исходного балла для комплексной оценки комбинированного знания (например, одновременного знания лексики и грамматики иностранного языка) приводит к существенному искажению результата, поскольку распределение итогового тестового балла становится резко несимметричным (правосторонним). Выходом из положения является нелинейное преобразование исходного балла (извлечение квадратного корня при одновременной двухаспектной проверке). В дальнейшем с целью удобства это значение может подвергаться линейному преобразованию.

Пример: Итоговый тест для проверки знаний лексики и грамматики студентов 5-го курса педагогического вуза по дисциплине «немецкий язык» состоит из 64 заданий, при этом задания стратифицированы по необходимому для их выполнения уровню знаний грамматики и лексики. Итоговая оценка (на 100-балльной шкале) может выводиться из соотношения $T_{ит} = 12,5\sqrt{T_{исх}}$, где $T_{исх}$ – исходный балл (от 0 до 64). Нормирующий множитель 12,5 обеспечивает вариацию итоговой оценки в пределах 0–100 баллов при заданном количестве заданий в тесте. При отсутствии статистических данных об уровне трудностей заданий можно воспользоваться делением диапазонов грамматических и лексических трудностей заданий на 8 уровней в каждом диапазоне, подключая к каждому уровню лексику и грамматику, изучаемую в соответствующем семестре. После проведения предтестирования задания перераспределяются на основе набранной статистики.

При необходимости простого ранжирования испытуемых необходимость в нелинейном преобразовании исходного тестового балла отсутствует, однако при необходимости включения тестового балла в некоторые необходимые для целей управления учебным процессом конструкты (например, для формирования рейтинга студента) лишь нелинейное преобразование исходного тестового балла по описанной схеме гарантирует корректность дальнейших процедур оценивания и принятия решений.